Structure matters

Citation for published version:
Waters, Z, Kovanovic, V, Kitto, K & Gasevic, D 2015, Structure matters: Adoption of structured classification approach in the context of cognitive presence classification. in *Proceedings of the 11th Asia Information Retrieval Society Conference.* Lecture Notes in Computer Science, Springer, pp. 227-238, 11th Asia Information Retrieval Societies Conference, AIRS 2015, Brisbane, Australia, 2/12/15. https://doi.org/10.1007/978-3-319-28940-3_18

Digital Object Identifier (DOI):
10.1007/978-3-319-28940-3_18

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 11th Asia Information Retrieval Society Conference

Structure matters: Adoption of structured classification approach in the context of cognitive presence classification

No Author Given

No Institute Given

Abstract. Within online learning communities, receiving timely and meaningful insights into the quality of learning activities is an important part of effective educational experience. Commonly adopted methods – such as the Community of Inquiry framework – rely on manual coding of online discussion transcripts, which is a costly and time consuming process. There are several efforts underway to enable the automated classification of online discussion messages using supervised machine learning, which would enable the real-time analysis of interactions occurring within online learning communities. This paper investigates the importance of incorporating features that utilise the structure of online discussions for the classification of "cognitive presence" – the central dimension of the Community of Inquiry framework focusing on the quality of students' critical thinking within online learning communities. We implemented a Conditional Random Field classification solution, which incorporates structural features, which may be useful in increasing classification performance over other implementations. Our approach leads to an improvement in classification accuracy of 5.8% over current existing techniques when tested on the same dataset, with a precision and recall of 0.630 and 0.504 respectively.

Keywords: Text Classification, Conditional Random Fields, Online Learning, Online Discussions

1 Introduction

The classification of social interactions occurring among individuals who participate in an online community is an important research problem. Not all participant contributions have the same value, with some being more thoughtful than others. This problem is particularly important in an educational domain, where online discussions are often being used to support both fully online and blended models of learning [7]. There is a substantial body of research that aims to produce online learning communities that foster higher-order learning and thinking among students. One prominent framework for approaching this problem is the Community of Inquiry (CoI) model [8] which describes the important dimensions of learning in online communities, and provides a quantitative coding scheme for their assessment.

Despite wide adoption by online education researchers, coding online discussion transcripts according to the CoI schemas is a manual and labor-intensive task, often requiring several coders to dedicate significant amounts of time to code each of the discussion messages. This approach i) does not enable for a real-time feedback on the quality of learning interactions, and ii) limits the wider adoption of the CoI framework by educational practitioners. This problem makes the task an ideal candidate for automation, and a number of approaches aimed at automating the process of coding transcripts using machine learning techniques are in development [22, 2, 17]. While these approaches have produced promising results, their text classification models currently make class predictions on a per-message basis, using only features derived from a single post, without consideration of the context of a post or of the preceding classification sequence. Given that human coders take discussion context into account during the classification process, and that the underlying construct of cognitive presence develops over time [9, 7], it seems likely that structural classification features can be used to model context in a similar fashion, and that these might improve classification accuracy.

This paper presents the preliminary results of a new approach to the automated analysis of online discussions within online learning communities using Conditional Random Fields (CRFs) [27], which is a novel contribution towards the automatic text-classification of online discussions using the CoI framework. Our results show that the use of structural features in a CRF model produce a higher classification accuracy than currently available methods. In section 2, the CoI model is briefly introduced, with an emphasis upon examining current approaches to analysing community participants' "cognitive presence". Related applications of CRFs to online discussions are also reviewed. Section 3 outlines an experiment that aims to improve on existing approaches by combining structural features with a Linear-Chain CRF model. The results of this experiment are presented in section 4, where they are compared against current approaches and human accuracies. Structural features and their potential use across a number of contexts and discussion media are discussed in section 5. Following this, the limitations of the experiment are discussed, which form the basis of the future work directions (section 5.1). Finally, the research presented in this work is summarised in section 6.

2 Background Work

2.1 The Community of Inquiry (CoI) framework

Overview. The Community of Inquiry (CoI) framework [8, 7] proposes three important dimensions (presences) of inquiry-based online learning:

1. **Teaching presence** defines the role of instructors before and for the duration of a course, consisting of i) direct instruction, ii) course facilitation, and iii) course organization and design.

2. **Social presence** provides insights into the social climate between course participants. It consists of i) affective communication, ii) group cohesion, and iii) interactivity of communication.
3. **Cognitive presence** is a central component of the framework and defines phases in the development of cognitive and deep thinking skills in online learning community [8].

The CoI framework defines multi-dimensional content analysis schemes [4] for the coding of student discussion messages, which is the main unit of analysis used to assess the level of the three presences. This framework has gained considerable attention in the educational research community, with a large number of replication studies and empirical validations (cf. [10, 9]). Overall, the CoI framework and its coding schemes show sufficient levels of robustness (see section 3.1 for an example) resulting in widespread adoption of the framework in the online education research community [10].

Of particular interest is the level of cognitive presence exhibited by the community members, due to its indication of their critical thinking. It is defined as the *"extent to which the participants in any particular configuration of a community of inquiry are able to construct meaning through sustained communication."* [8, p11], and is operationalized through a practical inquiry model which defines the four phases of the inquiry process that occurs during learning [8]:

1. **Triggering**: In the first phase, students are faced with some problem or dilemma which triggers a learning cycle. This typically results in messages asking questions and expressing a sense of puzzlement.
2. **Exploration**: This phase is primarily characterized by the exploration – both individually and in group – of different ideas and solutions to the problem at hand. Brainstorming, questioning, leaping into conclusions, and information exchange are the primary activities in the exploration phase.
3. **Integration**: After exploring different ideas, students synthesize the relevant ideas which ultimately leads to construction of meaning [8]. From the perspective of an instructor, this is the most difficult phase to detect as integration of ideas is often not clearly visible in discussion transcripts.
4. **Resolution**: In the final phase, students apply the newly constructed knowledge to the original problem, typically in the form of hypothesis testing or the building of a consensus.

Challenges of CoI framework adoption. One of the biggest practical challenges in adoption of the CoI framework – and other transcript analysis methods – is that it requires experienced coders and substantial labor-intensive work to code discussion messages for the levels of three presences [17, 4]. As such, it is argued that this and similar approaches have had very little practical impact upon current educational practices [4]. More recently, Arbaugh et al. [1] developed a 34-item Likert-scale survey instrument which can be used to measure levels of the three CoI presences more easily. However, there are still many challenges common for this type of instruments (i.e., self-selection of survey participants, low response rates, post-course administration) that makes the survey instrument usable primarily for post-course evaluation and research purposes rather than for in-class use and real-time interventions. To enable for a more proactive use of the Community of Inquiry framework by the course instructors, there is a need for an automated content analysis of online discussions that would provide instructors with a real-time feedback about student learning activities [15].

2.2 Automated classification of student discussion messages

Using machine learning to classify student messages in online discussions is generally a challenging task. Kovanović et al. [17] presented an approach to classifying cognitive presence from online discussions, using a Support Vector Machine (SVM) classification model, which achieved classification accuracy of 58.84%. While the results of this work are promising, they were achieved using only lexical features derived from the content of each individual discussion message. These features consisted of various N-grams, POS tags, name entity counts and dependency tuples, as well as intuitive features such as whether a post or reply is the first in a discussion thread. In contrast, human coders typically utilise contextual information – such as the structure of online discussions – when making their coding decisions. Because of this, it is worth investigating how structural features about a discussion may further improve classification performance.

Beyond the CoI framework, many studies have acknowledged that accounting for the relationships between individual messages and the latent structure of discussions may improve classification performance for transcript analysis [26, 5, 23]. Specifically, Ravi and Kim [23] suggests that using features derived from a previous message can be a positive indicator for classification of the next post along in a discussion. Other related work in threaded-discussion classification that seeks to incorporate the structural features of discussions is becoming increasingly common [6, 29, 14]. The most common type of structural features utilised include a post's position relative to others in a discussion, whether a post is the first or the last in a thread, how similar a post is as compared to its neighbours, and how many replies a post accrued. For this study, we attempt to account for the latent structure between posts in a discussion by incorporating these features into a Conditional Random Field approach.

2.3 Conditional Random Fields for Automated Detection of Cognitive Presence

We have implemented a Conditional Random Field (CRF) classification model [27] to annotate posts within a discussion with the phases of cognitive presence. Unlike traditional linear text classification methods, Conditional Random Fields consider the label sequence of a data set. Because of this, Conditional Random Fields have found numerous applications in natural language processing (NLP) tasks, such as part-of-speech (POS) tagging [18], document segmentation and summarisation [25], and gene prediction from biological sequence data [3].

Recent related research has extended CRFs to online forum discussions, where posts and interactions between participants are sequential in nature. Wang et al. [29] applied CRFs to discussion forums to learn the reply structure of forum interactions. This was achieved by using rich features that capture both short and long range dependencies within posts of an online discussion such as the lexical content similarity between two neighbouring posts. Similarly, FitzGerald et al. [6] combined the lexical features of posts with a Linear-Chain CRF to detect high quality comments in blog discussions, such as the word and sentence count of the post. Moreover, FitzGerald et al. [6] postulates that there exists sequential dependencies between posts in a forum, which emphasises the usefulness of structural features derived from the entire discussion, as well as lexical features from a single post. To date, CRF classification has not been applied to the problem of automating the detection of Cognitive Presence in online discussion transcripts. Here, we show that making this step improves the accuracy of classification when compared with the current best practices.

3 Methods

3.1 Dataset

The data used in this study comes from a six offerings of a fully-online masters-level research-oriented course in software engineering at a Canadian public university. This is the same dataset as was used in the study by Kovanović et al. [17] which makes for more accurate and direct comparison between the two different classification approaches. In total, the data consists of 1,747 messages produced by 81 students. Each message was coded by two experienced coders who achieved an excellent level of coding agreement of 0.97 Cohen's Kappa, with only 32 coding disagreements in total. Table 1 shows the distribution of messages in different phases of cognitive presence. The details of course structure and organization are explained in detail in Kovanović et al. [16], Gašević et al. [12].

Table 1. Cognitive Presence Coding

ID	Phase	Messages	(%)
0	Other (no signs of cognitive presence)	140	8.01%
1	Triggering Event	308	17.63%
2	Exploration	684	39.17%
3	Integration	508	29.08%
4	Resolution	107	6.12%
	All phases	1747	100%

3.2 Data Preprocessing

In this dataset, online discussions form a tree-like hierarchical structure (i.e., each discussion message can receive replies which can also receive replies). This presents a problem; in order to train and test our linear-chain CRF implementation, the structure of the data must be linear, as opposed to the current tree structure. In order to obtain appropriate sequences of data, sub-threads were extracted such that every sequence of posts from the root node to every leaf node in a tree was obtained. To obtain reliable results, these 1747 sub-threads must be remerged after classification to produce one classification per message in a discussion; this remerging process in described in section 4.1. While other CRF models will accept hierarchical structures (e.g., such as Tree-Structured and Hierarchical CRFs), our method of using a linear-chain model over other approaches due to the size constraints imposed by the dataset. This there are only 84 discussion threads in total to use for training and testing a tree-structured model, as opposed to a large number of message sequences used in our linear-model.

In addition to the extraction of linear sequences, the discussion threads in the data set were split into two sets; one for training and testing the CRF model, the other for validation from which our results are derived. These threads were split 70/30/10% for training, testing and validation, respectively.

3.3 Classifier Implementation

For this study, we implemented a Linear-Chain Conditional Random Field (LC-CRF) model to predict the phases of cognitive presence occurring in online discussions. This LCCRF was implemented in Java using the Mallet library [21], which is a widely used open source toolkit for machine learning. This library was extended as needed to suit our experimental requirements.

3.4 Classification Features

Many of the features used for the purpose of this study were extracted using the various functionalities of the Stanford CoreNLP Java library [20], and are derived from the related work in our literature review. Each post in the discussion is described by a feature vector that attempts to encapsulate both lexical and structural features. In addition to word unigrams, lexical features were derived from the text content of a post itself, and structural features were used to indicate where a post resides in the context of the entire discussion thread. These features are presented below:

1. **Entity Count** is the number of entities within a post as found by the Stanford CoreNLP Named Entity Recognition (NER) tool. The rationale behind using this feature is that discussion participants posting exploration comments are more likely to introduce a number of entities through their exploration of ideas.

2. **First Post** and **Last Post** are boolean features that are set to true when a post is the first and last in a discussion respectively. This feature represents the implicit structure of the discussion, where it is intuitive to believe that most Triggering phases occur at the start of a discussion.
3. **Comment Depth** is the number assigned to a post based on its chronological order within a discussion thread.
4. **Post Similarity** of the previous and next post in a discussion, this feature is calculated by obtaining the cosine similarity of two TF-IDF weighted vectors. The post similarity features assist in incorporating the local structure of the discussions, where it is expected that some phases of cognitive presence differ significantly from one another, and some only slightly.
5. **Word and Sentence counts** capture the number of words and sentences within a particular post. It is expected that when a discussion is reaching the integration and resolution phases, there is a lot more content due to the synthesis and integration of ideas.
6. **Number of Replies** to a post, which provides the classifier with the intuition that the earlier phases of cognitive presence (Triggering and Exploration) will have more replies than the later phases. Additionally, this feature also helps model the implicit structure within a discussion, giving the classifier an indication of how large the discussion is. The rationale behind this feature is that the triggering and exploration phases would generally have more replies than the integration and resolution phases.

These features form a feature vector for each message in a discussion thread. Because our classifier is sequential, these feature vectors are combined to form a feature vector sequence used in Mallet for training and testing our CRF classification model.

4 Results

The aim of this experiment is to investigate whether classifying posts in sequence, with the addition of structural features can improve upon the current approach of identifying cognitive presence in online learning discussions. A comparison between this experiment and the approach with the current highest accuracy is described in table 2.

Before remerging the discussion threads, the CRF model achieved an accuracy of 67.2%, and 0.515 and 0.620 precision and recall respectively and a F-measure of 0.562. Because sub-threads were extracted for this experiment (detailed in section 3.2), messages found earlier in the discussion threads have been classified multiple times. As a result of this, these accuracies are optimistically high due to multiple correct classifications diluting the overall classification accuracies. To overcome this drawback, the discussion threads were remerged back into their original hierarchical form in order to obtain reliable classification results.

Table 2. Comparison of Results

Approach	Cohen's Kappa	Accuracy
Kovanović et al. [17]	0.410	58.4%
LCCRF	0.482	64.2%

4.1 Re-merging Discussion Threads

As mentioned earlier in section 3.2, every message sequence from a root post to every leaf node in a discussion was extracted to produce an appropriate linear sequence to train the LCCRF. This means that the earlier posts in a discussion may have been classified multiple times. The the predicted phase need not necessarily be the same for these multiple classifications; a post that was classified as Triggering in one sequence might be classified as Exploration in the next sequence it appears in. In order to obtain one classification result for each message in a threaded discussion, the sub-threads were remerged using a majority vote mechanism. This method of remerging posts results in a final accuracy of 64.2%, obtained for the validation set. A large majority of posts that were classified multiple times belonged to the Triggering label, but many of these multiple classifications were correctly identified. Thus, the resulting drop in performance is representative of the general classification accuracy obtained by the linear CRF. Overall, the CRF implementation appears to perform well at this type of classification task, with an overall precision and recall of 0.630 and 0.504 respectively and a F-measure of 0.559. Moreover, our implementation achieve a Cohen's Kappa value of 0.482, which a widely used metric for judging the overall performance and reliability of a coding or categorisation scheme. Moreover, obtaining a Cohen's Kappa value that is similar to that of human coders is the underlying objective of there experiment, improving upon this value forms a significant role in our future work.

5 Discussion

Our CRF approach shows a promise for the classification of cognitive presence from discussions within an online learning community. Moreover, the results of this work show a modest improvement over the work conducted by Kovanović et al. [17], who presented an accuracy of 58.4% as seen in table 2. The key differences in these two approaches is clear: our approach considers discussion messages in sequence modelled the CRF, utilising features that attempt to convey the context of the discussion; whereas the work presented by Kovanović et al. [17] considers each message separately, relying on primarily lexical features using a SVM.

Our results show that our CRF utilising structural features is well suited to this text classification task. Using this approach, the classifier appropriately models the dependencies between messages in online discussions. This feature-set allows for a contrast between posts that would otherwise contain very similar lexical features. By combining these features, the probabilistic CRF implementation is better at modelling the dependencies between posts, leading to increased predictive performance. This improvement demonstrates a preliminary evidence of the importance of modelling the structure of discussions and their posts to improve the automation of cognitive presence detections. Because of the nature of structural features, they may also be useful for training a classification models across discussion platforms that share the same threaded messages. However, further research is required to support this claim. Moreover, this model was trained using a dataset obtained from one online course. Thus, future research needs to consider data sets from courses in other subject areas and delivery mode (i.e., blended learning).

As seen in Table 1, the distribution of phases (class labels) is largely uneven. This disparity between the individual phases of cognitive presence is evidence in the predictive performance of our classifier, where the predictive performance with respect to the lowest represented phases is typically less than that of their higher represented counterparts. This is a commonly acknowledged aspect of the type of collaboration within online learning communities, where learners typically do not progress to the resolution phase of cognitive presence Garrison et al. [11], Gašević et al. [12]. Future attempts at automation may benefit from a method of accounting for this uneven distribution of class labels.

5.1 Limitations and future work

There are a some limitations to this study which will be addressed in future work.

One key limitation of this work is contextual, our results may be biased towards the single course from which the dataset was derived. Moreover, there are a number of different platforms in which online learning discussions can take place. For example, a learning community using Social Media may be more informal in nature than one conducted in an institutes formal discussion forum. Using a model trained on one community may not produce reliable results for another community. This problem may be overcome with the usage of structural features, and further validation on datasets from other communities will be required. Compiling future datasets is a particular problem for this type of research, and is one that will be addressed in future work.

Other approaches to move towards automating the coding process will be investigated as future work. Because this approach uses a linear-chain model, some dependencies between messages in an online discussion may be missed. However, this linear model allows for the implementation of coding practice rules used by various CoI coding schemes, such as "coding up" – i.e., when a message has traces of two phases of cognitive presence, it is coded with the higher phase[16]. Despite this, approaches that might better model dependencies hierarchical structures, such as a tree CRF may further improve on our current accuracy.

In order to replace the current approach to analysing online learning communities with manual hand-coding transcripts, we aim to achieve Cohen's Kappa value of close to 0.80, which indicates an almost perfect agreement among coders according to Landis and Koch's Landis and Koch [19] interpretation of Cohen's Kappa. Our CRF approach achieved a Kappa value of 0.482, which indicates a moderate agreement according to Landis and Koch, but will require further improvement before machine learning techniques can replace hand coders. Future work will aim to further improve our classifier's performance. Specifically, we plan to further improve our model by: (i) evaluating our model on another, larger dataset with a more even distribution of phases; (ii) seeking additional features that may improve upon our current accuracies, such as Coh-Metrix Graesser et al. [13] and features derived from the Linguistic Inquiry and Word Count (LIWC) framework Tausczik and Pennebaker [28] that are commonly used to characterise cognitive processing associated with comprehending and producing text and discourse, and; (iii) better modelling the dependencies between threaded discussions using a Tree-Structured CRF model approach

6 Conclusion

In this work, we presented a new approach to automating the detection of the four phases of cognitive presence arising in online discussions. By reconceptualising online discussions as a sequence prediction problem, we predicted a sequence of labels (i.e. the phases of cognitive presence) for a sequence of messages. This allowed us to use a linear chain Conditional Random Field model for classification, which incorporates structural features of online discussions rather than just the lexical features that have previously been applied to solving this problem. This approach to automating the detection of cognitive presence has shown promise, with moderate improvements over alternative approaches with an accuracy of 64.2% and a Cohen's Kappa value of 0.482. However, classification accuracies are not yet high enough to replace the current approach of manually coding transcripts. Further improving this model is a priority for future work where we aim to further evaluate the model on alternative datasets, investigate additional features, and attempt to better model the dependencies between posts using a tree-structured CRF model.

References

[1] Arbaugh, J., Cleveland-Innes, M., Diaz, S.R., Garrison, D.R., Ice, P., Richardson, J.C., Swan, K.P.: Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry framework using a multi-institutional sample. The Internet and Higher Education 11(3–4), 133–136 (2008)

[2] Corich, S., Hunt, K., Hunt, L.: Computerised Content Analysis for Measuring Critical Thinking within Discussion Forums. Journal of e-Learning and Knowledge Society 2(1), 47 – 60 (2012)

[3] Culotta, A., Kulp, D., McCallum, A.: Gene prediction with conditional random fields (2005)

[4] Donnelly, R., Gardner, J.: Content analysis of computer conferencing transcripts. Interactive Learning Environments 19(4), 303–315 (2011)
[5] Feng, D., Shaw, E., Kim, J., Hovy, E.: An intelligent discussion-bot for answering student queries in threaded discussions. In: Proceedings of the 11th international conference on Intelligent user interfaces. pp. 171–177. ACM (2006)
[6] FitzGerald, N., Carenini, G., Murray, G., Joty, S.: Exploiting conversational features to detect high-quality blog comments. In: Advances in Artificial Intelligence, pp. 122–127. Springer (2011)
[7] Garrison, D.R.: Thinking Collaboratively: Learning in a Community of Inquiry. Routledge, New York, NY (2015)
[8] Garrison, D.R., Anderson, T., Archer, W.: Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education. American Journal of Distance Education 15(1), 7–23 (2001)
[9] Garrison, D.R., Anderson, T., Archer, W.: The first decade of the community of inquiry framework: A retrospective. The Internet and Higher Education 13(1–2), 5–9 (2010)
[10] Garrison, D.R., Arbaugh, J.: Researching the community of inquiry framework: Review, issues, and future directions. The Internet and Higher Education 10(3), 157–172 (Jan 2007)
[11] Garrison, R., Cleveland-Innes, M., Fung, T.S.: Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. The Internet and Higher Education 13(1–2), 31–36 (2010)
[12] Gašević, D., Adesope, O., Joksimović, S., Kovanović, V.: Externally-facilitated regulation scaffolding and role assignment to develop cognitive presence in asynchronous online discussions. The Internet and Higher Education 24, 53–65 (2015)
[13] Graesser, A.C., McNamara, D.S., Kulikowich, J.M.: Coh-metrix providing multilevel analyses of text characteristics. Educational Researcher 40(5), 223–234 (2011)
[14] Jin, W.: Blog comments classification using tree structured conditional random fields. Ph.D. thesis, University of British Columbia (Vancouver (2012)
[15] Kovanovic, V., Gasevic, D., Hatala, M.: Learning Analytics for Communities of Inquiry. Journal of Learning Analytics 1(3), 195–198 (2014)
[16] Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., Adesope, O.: Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. The Internet and Higher Education 27, 74–89 (2015)
[17] Kovanović, V., Joksimović, S., Gašević, D., Hatala, M.: Automated Content Analysis of Online Discussion Transcripts. In: Proceedings of the Workshops at the LAK 2014 Conference co-located with 4th International Conference on Learning Analytics and Knowledge (LAK 2014). Indianapolis, IN (Mar 2014), http://ceur-ws.org/Vol-1137/
[18] Lafferty, J., McCallum, A., Pereira, F.C.N.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data (2001)

[19] Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. biometrics pp. 159–174 (1977)

[20] Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. pp. 55–60 (2014), http://www.aclweb.org/anthology/P/P14/P14-5010

[21] McCallum, A.K.: Mallet: A machine learning for language toolkit. Tech. rep. (2002), http://mallet.cs.umass.edu

[22] McKlin, T., Harmon, S., Evans, W., Jones, M.: Cognitive presence in web-based learning: A content analysis of students' online discussions. In: IT Forum. vol. 60 (2002)

[23] Ravi, S., Kim, J.: Profiling student interactions in threaded discussions with speech act classifiers. Frontiers in Artificial Intelligence and Applications 158, 357 (2007)

[24] Rourke, L., Anderson, T.: Validity in quantitative content analysis. Educational Technology Research and Development 52(1), 5–18 (2004)

[25] Shen, D., Sun, J.T., Li, H., Yang, Q., Chen, Z.: Document summarization using conditional random fields. In: IJCAI International Joint Conference on Artificial Intelligence. pp. 2862–2867 (2007)

[26] Soller, A., Lesgold, A.: A computational approach to analyzing online knowledge sharing interaction. In: Proceedings of Artificial Intelligence in Education. pp. 253–260 (2003)

[27] Sutton, C., McCallum, A.: s. Foundations and Trends in Machine Learning 4(4), 267–373 (2011)

[28] Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: Liwc and computerized text analysis methods. Journal of language and social psychology 29(1), 24–54 (2010)

[29] Wang, H., Wang, C., Zhai, C.X., Han, J.: Learning online discussion structures by conditional random fields. In: SIGIR'11 - Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 435–444. Beijing (2011)